CLAIMS

1.   In an information retrieval system having a user, a current query, a search engine that is responsive to the current query to retrieve data generally relevant to the current query from a data collection, and data generally relevant to the current query retrieved from a target data collection, a method for refining the current query, the method comprising:

indexing a subset of the retrieved data into a vector space;

receiving information regarding the relevancy of a subset of the retrieved data;

ranking a subset of the retrieved data in accordance with the relevancy information;

forming at least one hypothesis regarding modifications to the current query, each hypothesis based at least in part on one or more of:

the rankings and

analysis of locations of the data in the vector space;

creating at least one candidate query in accordance with at least one of the formed at least one hypothesis;

applying at least one of the at least one candidate queries to the retrieved data, employing a search engine responsive to the at least one applied candidate query to retrieve data relevant to the at least one applied candidate query;

ranking the data retrieved in response to the applied candidate query in accordance with the received relevancy information;

comparing the ranking of data retrieved in response to the applied candidate query and the ranking of data retrieved in response to the current query with the received relevancy information

choosing the query which produces results more closely matching the ranking of the retrieved data as a refined query

thereby refining the current query.

2.   A method for forming a hypothesis for refining a query, the method comprising:

indexing a set of data into a vector space;

receiving information regarding the relevancy, with respect to information needs, of a subset of the data;

forming at least one hypothesis regarding modifications to a query based at least in part on analysis of:

location of the data in the vector space, and

a subset of the received relevancy information.

3.   A method for testing a hypothesis for refining a current query, the method comprising:

indexing a set of data into a vector space

receiving information regarding the relevancy, with respect to information needs, of a subset of the data;

ranking a subset of the data received in response to a current query, based on the degree of responsiveness to the relevancy information;

modifying the current query to incorporate modifications derived from a hypothesis for refining the current query;

applying the modified query to a subset of the data;

ranking a subset of the data retrieved in response to the modified query based at least in part on:

location of the data in the vector space, and

a subset of the received relevancy information;

determining

which of:

the ranking of the subset of data retrieved in response to the current query and

the ranking of the subset of data retrieved in response to the modified query

better corresponds to the relevancy information

thereby testing the hypothesis that incorporation of the modifications improves the correspondence between the query and the received relevancy information.

4.   A method for forming a hypothesis for refining a strategy for ranking the results of a data retrieval operation, the method comprising:

indexing a set of data into a vector space;

receiving information regarding the relevancy, with respect to information needs, of a subset of the data;

forming at least one hypothesis regarding modifications to a ranking strategy based at least in part on analysis of:

location of the data in the vector space, and

a subset of the received relevancy information.

5.   A method for testing a hypothesis for refining a strategy for ranking the results of a data retrieval operation, the method comprising:

indexing a set of data into a vector space;

receiving information regarding the relevancy, with respect to information needs, of a subset of the data;

ranking a subset of the data based on a current ranking strategy;

modifying the ranking strategy to incorporate at least one modification based at least in part on:

location of the data in the vector space, and

a subset of the relevancy information;

ranking a subset of the retrieved data based on the modified ranking strategy

determining which of:

the ranking of the subset of retrieved data based on the current ranking strategy and

the ranking of the subset of retrieved data based on the modified ranking strategy

better corresponds to the relevancy information

thereby testing the hypothesis that incorporation of the at least one modification into the current ranking strategy improves the correspondence between the ranking and the received relevancy information.

29